PROCESS FOR THE DEVELOPMENT OF IMAGE QUALITY METRICS FOR UNDERWATER ELECTRO-OPTIC SENSORS

James S. Taylor, Jr.[1], Brett Cordes[2], Sam Osofsky[3], and Anna Domnich[3]

Abstract-- **Electro-optic identification (EOID) sensors have been demonstrated as an important tool in the identification of bottom sea mines and are transitioning to the fleet. These sensors produce two and three-dimensional images that will be used by operators to make the all-important decision regarding use of neutralization systems against sonar contacts classified as mine-like. The quality of EOID images produced can vary dramatically depending on system design, operating parameters, and ocean environment, necessitating the need for a common scale of image quality or interpretability as a basic measure of the information content of the output images and the expected performance that they provide. Two candidate approaches have been identified for the development of an image quality metric. The first approach is the development of a modified National Imagery Interpretability Rating Scale (NIIRS) based on the EOID tasks. Coupled with this new scale would be a modified form of the General Image Quality Equation (GIQE) to provide a bridge from the system parameters to the NIIRS scale. The other approach is based on the Target Acquisition Model (TAM) based on Johnson's criteria and a set of tasks. The following paper presents these two approaches along with an explanation of the application to the EOID problem.**

Index Terms—Image Quality, EOID, Probability of Identification, NIIRS, GIQE, Target Acquisition Model

I. INTRODUCTION

Identification of mine-like objects (MLOs) is a pressing Fleet need. During mine countermeasures (MCM) operations, sonar contacts are classified as mine-like if their signatures are sufficiently similar to known signatures of mines. Each contact classified as mine-like must be identified as a mine or not a mine. During MCM operations in littoral areas, tens or even hundreds of MLOs must be identified. This time consuming identification process is currently performed by the human eye -- Explosive Ordnance Disposal (EOD) divers or Remotely Oper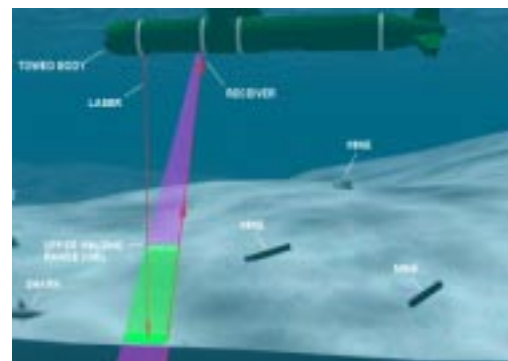ated Vehicles (ROVs) -- and is the rate limiting step in many MCM operations. A method to provide rapid visual identification of MLOs will dramatically speed up such operations. The method selected is electro-optic identification.

Two electro-optic identification (EOID) sensors are currently transitioning to the fleet. These are the Streak Tube Imaging Lidar (STIL), an addition to the AN/AQS-20/X and the WLD-1 (Remote Mine-hunting System) programs, and the Laser Line Scan (LLS), which will be part of the AN/AQS-14A(V1) program. Through these transitions, EOID will be a key element in implementation of Fleet plans for a robust organic MCM capability.

With these systems, the Fleet will have their first experience with high-resolution underwater electro-optical imagery. Therefore, it is necessary that a capability exists to measure the performance and effectiveness of these sensors under a given set of environmental conditions. Two approaches have been proposed. The first approach is the development of a modified image quality scale analogous to the National Imagery Interpretability Rating Scale (NIIRS) scale used by the surveillance and reconnaissance community. The other approach that applies is called Target Acquisition Model (TAM) and is based on a performance modeling technique developed, tested, and validated by the US Army. This paper will outline these two approaches and present the theory of each as well as the expected products in terms of MOPs and MOEs.

II. ELECTRO-OPTIC IDENTIFICATION SENSORS

The two laser identification systems set to transition to the Fleet are the Areté Associates Streak Tube Imaging LIDAR system in the AN/AQS-20 and WLD1, and the Northrop Grumman Laser Line Scan system in the AN/AQS-14A (V1). The laser line scan technology is illustrated in Figure 1.

			Form Approved OMB No. 0704-0188
Report Documentation Page		

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2002**	2. REPORT TYPE		3. DATES COVERED **00-00-2002 to 00-00-2002**
4. TITLE AND SUBTITLE **Process for the Development of Image Quality Metrics for Underwater Electro-Optic Sensors**			5a. CONTRACT NUMBER
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)			5d. PROJECT NUMBER
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Coastal Systems Station,6703 W. Highway 98,Panama City,FL,32407**			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES **Fifth International Symposium on Technology and Mine Problem, 22-25 Apr 2002, Monterey, CA**

14. ABSTRACT

Electro-optic identification (EOID) sensors have been demonstrated as an important tool in the identification of bottom sea mines and are transitioning to the fleet. These sensors produce two and three-dimensional images that will be used by operators to make the all-important decision regarding use of neutralization systems against sonar contacts classified as mine-like. The quality of EOID images produced can vary dramatically depending on system design, operating parameters, and ocean environment, necessitating the need for a common scale of image quality or interpretability as a basic measure of the information content of the output images and the expected performance that they provide. Two candidate approaches have been identified for the development of an image quality metric. The first approach is the development of a modified National Imagery Interpretability Rating Scale (NIIRS) based on the EOID tasks. Coupled with this new scale would be a modified form of the General Image Quality Equation (GIQE) to provide a bridge from the system parameters to the NIIRS scale. The other approach is based on the Target Acquisition Model (TAM) based on Johnson's criteria and a set of tasks. The following paper presents these two approaches along with an explanation of the application to the EOID problem.

15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Figure 1.   Illustration of the Laser Line Scan Technology.

The continuous wave laser illuminates a small moving spot on the ocean bottom.   A photomultiplier receiver, which is separated from the transmit beam, is synchronously scanned at the same rate to build up a raster-scanned image.   The resulting imagery is a 2 dimensional representation of the target of interest [1] as shown by the target in Figure 2.
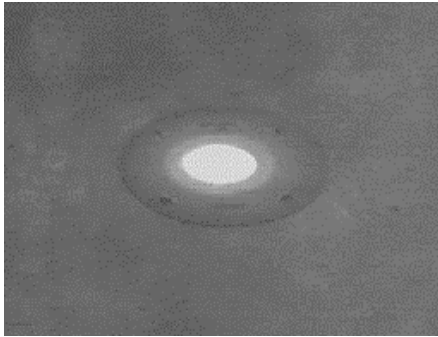

Figure 2. Laser Line Scan Image of a Truncated Cone.

The STIL technology was developed specifically for high-resolution three-dimensional imaging of underwater objects.[1] The STIL system is an active imaging system using a pulsed laser transmitter and a streak tube receiver to time resolve the backscattered light. The laser beam is diverged in one dimension using a cylindrical lens to form a fan beam. The backscattered light is imaged onto a slit in front of the streak tube photocathode by a conventional lens, and is time (range) resolved by electrostatic sweep within the streak tube, generating a 2-D range-azimuth image on each laser pulse. By orienting the fan beam perpendicular to the vehicle track, the along-track dimension is sampled by adjusting the Pulse Repetition Frequency (PRF) of the laser to the forward speed of the vehicle, thus sweeping out the three-dimensional ocean volume in a pushbroom fashion (Figure 3). [2]
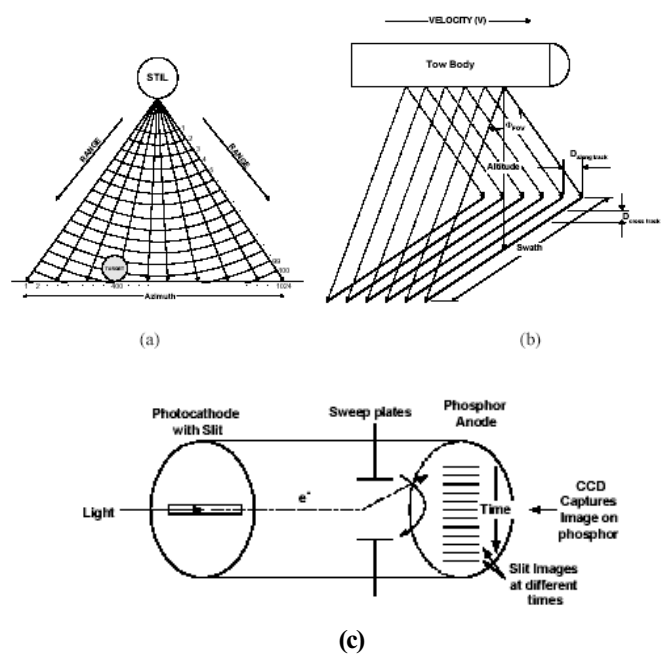

Figure 3.   Illustration of the Streak Tube Imaging Lidar (STIL). (a) and (b) Frontal view and Side view of the STIL showing STIL data collection with vehicle motion. (c) Streak tube receiver architecture. [2]

The precise temporal sampling of the STIL makes the sensor immune to ambient sunlight. The bottom return includes both time of flight information, which provides a quantitative measure of the height of the object above the bottom and the radiometric level that is proportional to the reflectivity of the bottom object.   Each laser shot thus provides range to and contrast of the bottom for each cross-track pixel. The imagery is rendered and the results are two images corresponding to contrast and range as shown by Figure 4. [2]
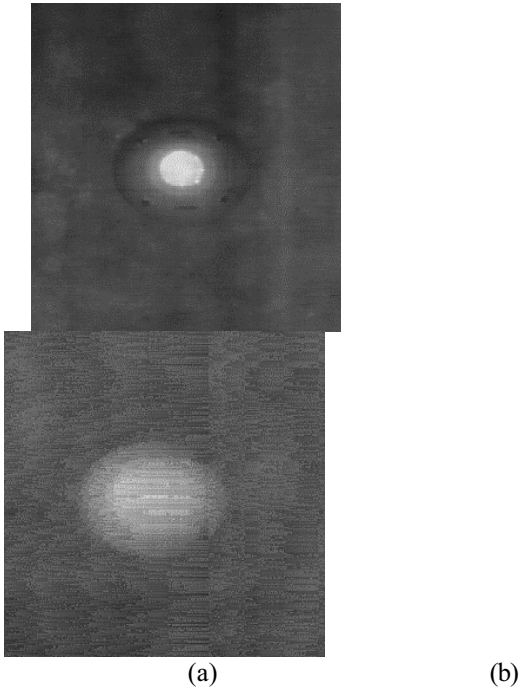
---

[1] Patent 5,467,122.

(a)   (b)

Figure 4. STIL (a) Contrast and (b) Range Imagery of a Truncated Cone.

Both Fleet systems have target cueing and snippet generation of EOID objects. Target cueing is basically a Region of Interest (ROI) matched filter algorithm that highlights certain areas of the image for the operator. Images snippets are generated either by the ROI algorithm or manually by the system operator.

Performance models are being developed so operations can be planned using water clarity measurements from oceanographic databases. The best water clarity data comes from sensors in the operation area during the operation. The STIL system is using a water clarity measurement to suggest an operational altitude during the sortie. Models have also been suggested to show reference targets as they would be seen through the current water conditions. Ideally, the warfighter would be able take whatever information is known of the environment, input that into a mission planning tool, and determine the expected performance, or P_{id}, to determine the utility of the EOID sensor for that mission on that day. Figure 5 illustrates this need.

Figure 5. Mission Planning Tools.

III. FLEET NEEDS

The two Measures of Effectiveness (MOEs) for MCM operations are time and residual risk. A given area can be said to be cleared of mines to a certain confidence level based on the search effort and the search results. A more confident declaration of a cleared area requires more time and assets to be applied in the operation. To calculate the MOEs, Measure of Performance (MOPs) must be established to quantify the individual system contribution. Some of these MOPs may seem familiar like Probability of Detection, others like Probability of Reacquisition (P_{reacq}) and Time to Identify (T_{id}) may not. All the MOPs must be defined in order to calculate the MOEs of time and residual risk. To get to the MOPs, it is necessary to relate the environmental variables back to sensor performance parameters. This constitutes the goal of this work.

In real world operations, the overall MCM MOE is sensitive to identification. Search systems provide large lists of mine-like contacts that will form the basis of prosecution lists for identification assets. Objects positively identified as mines can either change the ship operations area or call for neutralization assets to clear the mined area. The time required to identify the objects can be significant.

Both Fleet EOID systems will use sonars to help reacquire the initial contact increasing the P_{reacq}. The EOID systems will then be towed over the target for high resolution imaging of the object. The Probability of Identification (P_{id}) and Probability of False Identification (P_{fid}) are functions of water clarity and the altitude of the sensor over the target. Along with T_{id}, the time to identification against a single object, an analyst can calculate the time to perform the identification mission and the numbers of correctly and incorrectly identified mines. The MCM analysts would then have an estimate of the time and the residual risk to ships transiting in that area.

IV. APPROACHES

Performance prediction models have been developed for four classes of imaging systems: Synthetic Aperture Radar (SAR), downlooking visible and infrared (IR), battlefield target acquisition (low elevation angle visible and IR), and video. The primary performance metrics are NIIRS, probability of task performance, and subjective quality ratings.[4],[5] The primary purpose of an EOID performance metric is to give the operator a measure of the ability to distinguish between mines and clutter. In what follows, we will refer to this task as "identification."

Two of these approaches have been investigated as possible quality metrics for the EOID imagery. The first approach has been presented in recent EOID literature[7],[8] and centers on the development of an image quality scale based on the National Image Interpretability Rating Scale (NIIRS) that was developed by the surveillance and reconnaissance community to measure image quality of hard copy imagery. A logical extension to this new scale is a modified version of the General Image Quality Equation (GIQE) that was developed later to convert system design parameters to a NIIRS number.

The second approach known as the Target Acquisition Model (TAM) is based on the sensor modeling that the US Army has been doing for many years. It is based on the Johnson's criteria and Minimum Resolvable Contrast (MRC) and has been used to characterize the performance of acquisition sensors. This process revolves around the premise that there is a minimum resolvable contrast that is a function of spatial frequency.

Both approaches will be discussed here and comparisons will be presented. It should be noted that neither method has been developed with 3-dimensional data.

A. National Imagery Interpretability Rating Scale (NIIRS)

NIIRS was developed in the 1970s by a team of government and contractors. The scale was developed in response to the inability to measure the interpretability of imagery given the simple measures of image quality. The goal of the development was to provide a scale that would communicate what information could and could not be extracted from an image. A large sample of interpretation tasks was rated by a group of Image Analysts (IAs). The tasks ranged from detection of large facilities to the identification of small details, all of which had intelligence or military value. Various images of known quality were presented to another group of IAs for a determination of the most difficult task that could "just be accomplished."[4],[5],[6]

The resulting NIIRS scale was defined in terms of resolution with each level corresponded to a doubling or halving of resolution. The tasks were grouped into 10 levels (0-9) and were defined for the five military orders of battle (air, ground, naval, missile, and electronics). These criteria defined certain levels of interpretation.[4] For example, a NIIRS value of 2 will allow the analyst to detect large hangars at airfields while a NIIRS of 9 will allow the analyst to detect individual spikes on railroad ties.

The General Image Quality Equation (GIQE) was developed in the 1980s but not released until 1994.[4] The GIQE had been developed to relate system design parameters to the NIIRS scale. This was done because the next generation of surveillance and reconnaissance (S&R) sensors were specified in terms of NIIRS performance. The GIQE relates resolution, sharpness, and signal-to-noise ratio to a NIIRS number. The GIQE for visible EO is

$$NIIRS = 10.251 + a\log_{10} GSD_{GM} + b\log_{10} RER_{GM} \tag{1}$$
$$+ 0.656 H_{GM} - 0.344 (G/SNR)$$

where GSD_{GM} is the geometric mean ground sampled distance (in inches), RER_{GM} is the geometric mean of the normalized relative edge response, H_{GM} is the geometric mean height of overshoot due to MTF Compensation (MTFC), G = noise gain due to MTFC, SNR = signal-to-noise ratio, a = 3.32 if $RER \geq 0.9$ and 3.16 if $RER < 0.9$, and b = 1.559 if $RER \geq 0.9$ and 2.817 if $RER < 0.9$.[4],[5] The GIQE in its present form is not directly applicable to the EOID imagery. Substitution of the EOID-type numbers into the equation results in numbers that appear outside of the 0-9 scale, and even if the GSD were rescaled to a more relevant spatial scale range, the limits of the regression data set [5] used for deriving eq. (1) are significantly exceeded in SNR and RER in the EOID context.

B. Target Acquisition Model (TAM)

The TAM predicts the discrimination performance of operators who make their decisions based on examining a single grayscale image produced by an imaging sensor.

For a sensor which is well approximated as a linear, shift-invariant (LSI) imaging system, the TAM theory provides a metric for the quality of an image of a target that takes into account blurring, stochastic noise processes, and undersampling. The metric is the number N of Johnson bar cycles (with contrast equal to the target's characteristic contrast C) that can be discerned across the characteristic dimension D of the target by a human observer using the sensor (Figure 6).[4]
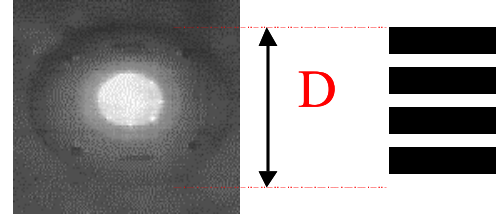

Figure 6. Illustration of Characteristic Dimension D and a 4-barTarget.

The method for computing N as a function of C, D and the degradation of the image resulting from scattering in the intervening medium constitutes much of the formalism of the TAM and is the application of the sensor models that are now being evaluated by the ONR EOID Research Program.

In the TAM, the relationship between N and probability of identification (P_{id}) must be established empirically for any given identification task and sensor. However, this calibration is embodied in a single number, N_{50}, which is the value of N for which the ensemble of human observers in the experiment is able to perform the identification task with 50% probability. Whatever formula gives P_{id} as a function of N and N_{50} can be expected to be a function only of the ratio N/N_{50} (under the assumption that *e.g.* doubling the size of the target should not effect P_{id} if the scale to which the sensor/operator can resolve is also doubled).

The Target Transfer Probability Function (TTPF) is the result of several empirical experiments and gives the probability of discrimination as a function of N/N_{50}. The values for the TTPF are given in Table 1.[1]

Table 1. Empirical values measured for the TTPF

Probability of discrimination	N/N_{50}
1.00	3.0
0.95	2.0
0.80	1.5
0.50	1.0
0.30	0.75
0.10	0.50
0.02	0.25
0	0

A formula has been empirically fit to this data [1] and is itself often referred to as the TTPF:

$$\tag{2}$$

It is important to note that determining N_{50} is a much easier task (that is, involving a smaller number of operator-decision experiments) than determining the full TTPF would be.

1) Application to 2D

The application of the TAM to the standard 2 dimensional data received from the EOID sensors is fairly straightforward. The block diagram in Figure 7 shows where the physics-based imaging models and the environmental parameters fit into the TAM.
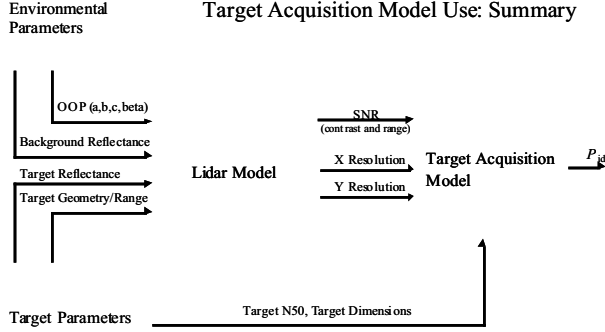

Figure 7. Application of TAM to EOID imagery.

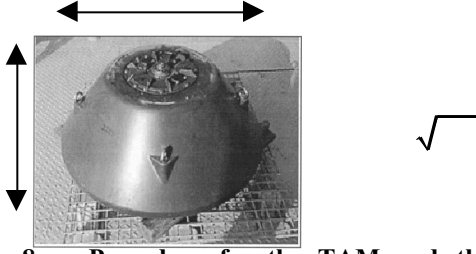The general procedure that would be used to apply the TAM to the EOID problem is shown in Figure 8.

Figure 8a. Procedure for the TAM and the EOID problem. The target's critical dimension d_c and characteristic contrast C are determined.
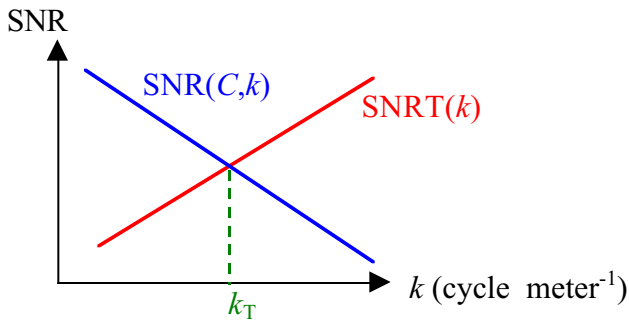

Figure 8b. The results of pre-existing, controlled experiments (human operators vs. patterns of varying spatial frequencies and contrasts on CRT screens) are combined with a model of the imaging process to derive an SNR Threshold (SNRT) for the Human Visual System (HVS). The objective is to discern Johnson bar patterns as a function of their spatial frequency k in the object plane. SNRT(k) is compared with SNR(C,k), the SNR calculated for HVS observations of Johnson bar patterns with spatial frequency k, where C, the contrast, is set equal to the characteristic contrast of the targets in the

target set. The two curves meet at a threshold spatial frequency k_T above which the bar patterns can no longer be perceived.
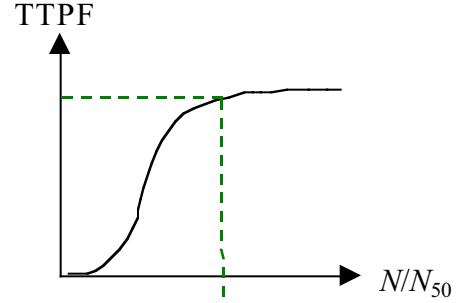

Figure 8c. The number N of just-discernible Johnson bar cycles, across the target critical dimension, is calculated from k_T, and the probability of performing the discrimination task is calculated using the TTPF.
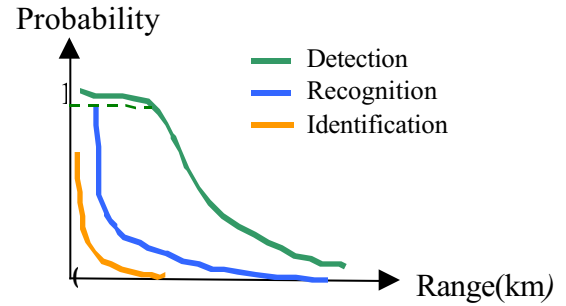

Figure 8d. N_{50} will vary with the type of discrimination task. In the US Army context, the three different discrimination tasks lead to three different N_{50}'s, and therefore three different curves for probability of performing a task as a function of e.g. range to target.

The application of this procedure on each of the 2D images (*e.g.*, LLS contrast, STIL contrast, and STIL range) should result in a P_{id} curve versus range for the 2-dimensional imagery. As a check, a regression analysis can be performed on the results from the 2D testing as a validation point on the TTPF. It should be noted that the TTPF is an ensemble probability that is applicable for the whole target set although performance against individual targets may be better or worse than the TTPF predictions.

2) Extension to 3D

While both sensors transitioning to the Fleet provide a contrast image, the STIL also produces a range map of the

ocean bottom as shown earlier. At the present time, there does not exist a methodology for the incorporation of a third dimension into the discrimination task for the Target Acquisition Model. Neither does there exist a quantification of the improvement that the extra information provides to the operator. As a part of this study, we have formulated an extension of the TAM process for the inclusion of the range information. This should provide the first look at quantifying the "range improvement."

For this 3D experiment, operators will be presented with two grayscale 2-D images of the ground truth, one indicating contrast and one indicating range.[2]

Just as a contrast image of the target has its N-value N_c, the associated range image of the target has an N-value N_r. Generally, these images will have to fill some 2-D region of (N_c, N_r)-space in order to encompass the images that a realistic system would encounter given varying altitudes, environmental conditions, and inherent target properties. As far as we know, there is no analog to the TTPF involving two N-valued images presented side-by-side. We will need to do operator testing on the (N_c, N_r)-space to determine the analog. We will consider various regression forms including interpolations of TTPF-type functions, but cannot predict in advance what functional form will best fit the data as, to our knowledge, generalization of the TAM to 2-image data has not been done before.

The region of (N_c, N_r)-space to be explored will be chosen by running existing physical models on part of the target set under environmental conditions ranging from estuarine through coastal and with low and high altitudes.

Although we cannot yet predict the best way to describe operator performance as a function of 2-image data quality, at least two hypotheses suggest themselves and will be considered once operator-testing data is measured. The benefit of hypotheses like these is that they restrict the regions on (N_c, N_r)-space for which operator testing must occur, and restrict the type of testing to determining some small number of N_{50}-like quantities. We should stress the fact that these hypotheses are merely to be checked; they may very well not be true and we are not relying on them to determine P_{id}.

a) Hypothesis 1: Statistically Independent Error Processes

Make the approximation that the processes inherent in misidentifying a contact given side-by-side images (in this case made with contrast and range data) are statistically independent. Furthermore, assume that the operator will use a "risk-averse" decision rule such that if he would choose to declare a mine present based on one type of image, then given both images he would still declare a mine present irrespective of whether he would have declared the second image alone to be mine or clutter.

Then the probabilities of interest can be written:

$$P_{RC}(\text{mine}|\text{mine}) \tag{3}$$

$$P_{RC}(\text{mine}|\text{clutter}), \tag{4}$$

where $P_{RC}(A|B)$ indicates the probability of declaring state A, conditioned on the true state being B, given both range and contrast images, and $P_{R,id}$ or $P_{C,id}$ indicate the probabilities of identification given only a range or contrast image respectively.

An important consequence of the assumptions for this hypothesis is that the P_{id}'s may vary significantly between clutter and mine elements of the test set, as can be seen in the differing expressions for eqs. (3) and (4). This would lead to an unavoidable dependence on the poorly known prior distributions of mine vs. clutter in the overall value for P_{id}. However, the hypothesis could lead to the easiest characterization of P_{id} on (N_c, N_r)-space because only two N_{50}-like calibrations would be required to characterize the whole space.

b) Hypothesis 2: Linearly Interpolated TTPF Regions

The line L of points where $N_c/N_{50,c} = N_r/N_{50,r}$, is potentially a special region of (N_c, N_r)-space because in a sequence of 2-image pairs along this line, the information quality in each of the images in a pair (at least as measured by the probability of discrimination given that image alone) is the same. The first assumption is that since the images in a pair have the same information quality, an image pair is acted on by the human operator as if it had come from a single type of imager, in the sense that the TTPF applies:

$$\text{TTPF}\left(\frac{N_c}{N_{50,c|rc}}\right), \tag{5}$$

where $N_{50,c|rc}$ is the value for N_c for which a 50% probability of identification obtains given pairs of images along L, and where TTPF(x) is given by equation (2). It should be noted that $N_{50,c|rc}$ differs from $N_{50,c}$ in that $N_{50,c|rc}$ is the value of N_c for which the users achieve 50% P_{id} given range/contrast pairs of images (along the line L of points where $N_c/N_{50,c} = N_r/N_{50,r}$) while $N_{50,c}$, is the value of N_c for which the users achieve 50% P_{id} given contrast images only. By definition of L, equation (5) could just as well have been written in terms of the analogous range-only quantities:

$$\text{TTPF}\left(\frac{N_r}{N_{50,r|rc}}\right), \tag{6}$$

where $N_{50,r|rc}$ is the value for N_r for which a 50% probability of identification obtains given pairs of images along L.

Although it may appear that two new parameters, $N_{50,c|rc}$ and $N_{50,r|rc}$, have been introduced, the fact that they are defined on L relates them and it is therefore useful to define a single new quantity $R_{50,rc}$ given by:

$$R_{50,rc} = \frac{N_{50,c|rc}}{N_{50,c}} = \frac{N_{50,r|rc}}{N_{50,r}}. \tag{7}$$

$R_{50,rc}$ is the ratio of N_{50} given the image pairs along L to the N_{50} given either range or contrast alone. The smaller $R_{50,rc}$ is measured to be, the more the data is indicating that the operator is able to take advantage of the side-by-side images. The unlikely possibility of a value larger than 1 would

---

[2]Other presentation algorithms, which may include combining the images into one image, may in fact work better. However, the first display to which the operator will be introduced is the side-by-side display currently planned for the AN/AQS-20/X console.

indicate confusion arising from the two images and suggest only one be used. The situation is indicated in Figure 9.
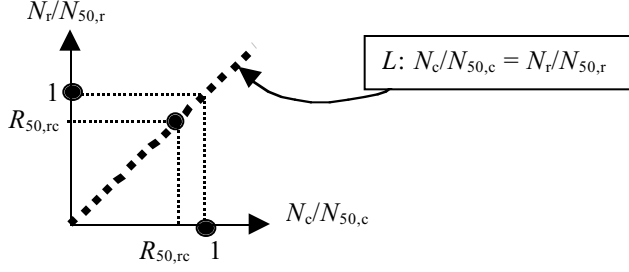

Figure 9. Special values in (N_c, N_r)-space. The thick dotted line indicates the identity set L defined by $N_c/N_{50,c} = N_r/N_{50,r}$, and the large black dots indicate points where $P_{id} = 50\%$.

Given the assumption about performance on L, it remains to extend the performance predictions to the rest of the space indicated in Figure 9. There is no obvious best way to do this. Perhaps the simplest suggestion is to approximate performance by a function that is constant along straight lines between points of equal probability where the probability has been measured -- on the identity and axes. Contours of constant probability resulting from this linear interpolation are indicated in Figure 10.
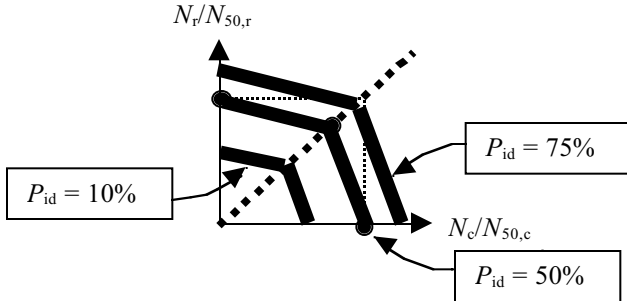

Figure 10. Sample of the contours of constant probability under the linear interpolation are indicated by the thick lines.

The formula giving P_{id} under these assumptions is:

$$\begin{array}{cccc} \underline{\hspace{1cm}} & \underline{\hspace{1.5cm}} & \underline{\hspace{1cm}} & \underline{\hspace{1cm}} \\ \\ \underline{\hspace{1cm}} & \underline{\hspace{1.5cm}} & \underline{\hspace{1cm}} & \underline{\hspace{1cm}} \end{array} \quad (8)$$

If the assumptions in this hypothesis hold, then only three N_{50}-like calibrations would be required to characterize the whole space.

V. NIIRS/GIQE VS. TAM AND THE "TAM-LIKE" APPROACH

Methods lying closer to TAM entail a much simpler operator-testing and regression problem then NIIRS/GIQE. Unlike the GIQE, TAM provides an SNR-based metric N, which already combines the quantitative image metrics such as

GSD, RER and SNR_{DC} rather than having to find regression coefficients for them from scratch. The TAM regression is on data ordered by this single number N to find a single number N_{50}, which characterizes the probability of performing a task. Extension of this method to the 2-image data will require additional regressions, whether using TAM or NIIRS/GIQE. Because of the multitude of unknowns associated with this extension, it would make sense to perform regression against quantities like GSD, RER and SNR_{DC}, in addition to regressing performance against N-values, and see which works best. But this is still a much more restricted problem than with the NIIRS/GIQE, because targets in the operator experiments (mines and clutter) can be restricted to have similar characteristic scales and contrasts (or range-contrasts), rather than necessarily studying a large range of spatial scales.

The approach we are describing has features mostly from TAM (limited variety in the operator test set, use of N-values, P_{id} as result), but also has features from the GIQE method (use of GSD, RER and SNR_{DC}) and a completely novel feature (2-image data). We will therefore call it a "TAM-like" approach.

In a sense, the TAM-like approach is more focused and less ambitious (giving higher probability of success) than NIIRS/GIQE, because like TAM it concentrates on identification among targets of similar size and inherent contrast. The products of this approach are actual probabilities, not just NIIRS-like numbers; these probabilities are in many cases more useful to the operator in the field.

Finally, the historical context that motivates the GIQE is not really relevant to EOID. GIQE was developed to predict NIIRS, a mature, pre-existing and widely-used metric. We would have to develop a new NIIRS and GIQE simultaneously. We conclude that a TAM-like approach such as we have described is the method of choice for the ONR EOID experiment.

VI. CONCLUSION

The NIIRS and TAM methods have both been studied for possible use of modeling MOEs and MOPs for EOID operations. Both approaches were deemed valuable, but with limitations. Subsequently, a "TAM-like" method, incorporating aspects of both metric models, has been developed and is a good candidate for well characterizing operator performance against EOID data. In particular, it appears to be considerably better for the ONR EOID experiment than a NIIRS/GIQE alone. Several hypotheses are available for simplifying the characterization of 2-image data (necessary because of the availability of both contrast and range data from STIL sensors).

The results of the ongoing development of this method are expected to be a viable, broadly applicable metric for EOID performance. Such a metric will fulfill the critical need for performance prediction when making essential tactical decisions in deploying EOID sensors in the fleet.

APPENDIX

For the US Army, decision-making tasks based on sensor images (called discrimination tasks) have been divided into three specific classes:
- "detection" (reasonable probability that blob is a tactical military vehicle),
- "recognition" (distinguish class e.g. truck, tank, APC)
- "identification" (object discrimination e.g. type of tank).

The Navy Mine Warfare community uses the following task definitions:

- Detection - Object/No Object (never really used, one system calls this "pre-detections")
- Classification - Mine-like Object/Non-Mine-like Object
- Identification - Mine/Non-Mine

REFERENCES

[1] M. P. Strand, "Underwater Electro-optic System for Mine Identification", Proceedings for Autonomous Vehicles in Mine Counter Measures Symposium, Monterey, CA, April, 1995

[2] A. Gleckler, Et Al, "Streak Tube Imaging Lidar for Electro-optic Identification", *Fourth International Symposium on Technology and the Mine Problem*, Monterey, CA, March, 2000

[3] G. C. Holst,, *CCD Arrays, Cameras, and Displays*: 2nd Edition, Bellingham, WA: SPIE Press, 1998.

[4] J. C. Leachtenauer and R.G. Driggers, *Surveillance and Reconnaissance Imaging Systems*, Norwood, MA: Artech House, 2001.

[5] J. C. Leachtenauer,. et al. "General Image-Quality Equation: GIQE." *Applied Optics* 36, pp. 8322-8 (1997).

[6] L. A. Maver, et al. "Imagery Interpretability Rating Scales," in *Digest of Technical Papers: International Symposium of the Society for Information Display* (Society for Information Display, Santa Ana, California, 1995), vol. **26**, pp. 117-120.

[7] M. P. Strand "Quantitative Evaluation of Environmental Noise In Underwater Electro-Optic Imaging Systems," Ocean Optics XIV.

[8] M. P. Strand, "Modeling the Performance on Underwater, Laser Based, Imaging Sensors," Ocean Optics XV.